Automating Cancer Registration

Challenges and Opportunities

Lynne Penberthy MD, MPH
IACR Meeting
September 17, 2007

Topics to be covered

- Current challenges in cancer registration
- Potential opportunities for cancer registration using electronic data and automation
- An example of cancer registration using electronic data and automation

Cancer Registration Challenges

- Data quality
- Data completeness
- Data overload - Information deficit
- Data consolidation dilemmas
  - Geographic dispersion of health care
  - Lengthy course of treatment

Challenge 1:*Data Quality*

Transcription and data entry errors

- Medical Record Number, personal identifiers, dates etc
  - May cause mismatch in case ascertainment details and follow up
  - May cause mismatch in consolidation at the central level

Challenge 2: ***Data Completeness***

Data from non-hospital locations for:

- Incident Cases
- Initial Treatment
- Follow Up
- missed without good consolidation and comprehensive reporting at the local and central level

Challenge 2: ***Data Completeness for Incident Cases***

Studies of incomplete reporting suggest that missed cases may represent biases in reporting[1-4]

- Rural, underserved populations
- Cases from *non-registry* hospitals
- Cases diagnosed and treated in physician offices

1. McClish DM and Penberthy LT.; 2005.
2. Penberthy LT, McClish DK, 2005.
3. McCLish, DK, Penberthy LT. "2004
4. Penberthy LT, McClish. 2003

Challenge 2: ***Data Completeness for Treatment***

- Cancer registries have limited sensitivity for chemotherapy capture (56-72%)[1-3]
- Studies suggest that missed treatment may represent biases based on reporting source and location
  - Increased use of chemotherapy in the outpatient or physician office setting

1. Malin 2002 (Breast)
2. Cress 2003 (Breast)
3. Du 2006 (Colorectal)

Challenge 2: ***Data Completeness for Treatment***

- Hormonal Therapy and Chemotherapy received as oral agents
  - Increased use of oral chemotherapy as mainstay for cancer treatment will raise additional challenges in chemotherapy capture
    - Likely purchased from pharmacy
    - No mandate to report (in US)

Challenge 3: *Data Overload*

Increasing data volume & information sources **within** a single institution require registrars to:

- Abstract information received ***at various times*** over the course of treatment
- Match & extract data from multiple reports to a single patient

Challenge 4: *Data Consolidation*

- Collating information from more than one location:
  - Inpatient facilities
  - Physician offices
  - Free standing radiation therapy centers
  - Free standing surgery centers
- Sharing patient information with colleagues at multiple institutions

How do these challenges impact cancer surveillance?

Biased & incomplete reporting of cases, treatments and outcomes

- Result in incomplete or erroneous case reports & statistics
- Raise questions on validity of analytic results

How do these challenges impact cancer surveillance?

Biased & incomplete reporting of cases, treatments and outcomes

- Limit the utility for outcomes assessments and research
  - Incomplete treatment for comparing differential outcomes
  - Incomplete outcomes ascertainment beyond survival
    - Recurrences, treatment of recurrences
- Limit the utility for clinical or public health purposes
  - Population comparisons such as rural, underserved or minorities data are difficult
  - Cancer survivorship

What can we do to maintain viability?

- Increase the value of registries
  - Enhance currency of data
  - Enhance completeness of information for clinical & public health use
- Improve cost benefit
  - Registries at both the hospital & central level are not revenue centers
  - Enhancing the efficiency or decreasing the cost will improve support

What can we do to maintain viability?

- To accomplish this formidable task I would propose that we need to rethink:
  - Data collection methods
  - Data elements we collect

A possible solution:
Automate and maximize the use of available electronic data

Automation possibilities: Improving efficiency & data quality

- Reduce ***direct data entry*** & ***transcription errors*** through:
  - *SELECTIVE* Automated upload and longitudinal capture of standardized patient specific data

Automation possibilities: Improving efficiency

Reduce the amount of paper handling required

Four months of paper pathology reports

Automation possibilities: Improving efficiency & security

- Make data exchange between registries more efficient, secure and complete
  - Sharing data electronically can be done in a secure manner
    - meeting electronic health data exchange standards such as HIPAA
  - Reducing the number of letters, telephone calls that are required for information gathering

Automation possibilities: Reducing data overload/Improving consolidation

- Automate the consolidation of information
  - from multiple data sources & locations
  - updating information ***over time*** to create longitudinal sets of records for each patient

Automation possibilities:
Increasing the value of registries

- Enhancing data completeness without adding to the registrars workload for:
  - Treatment
  - Recurrence & Subsequent Treatment
  - Comorbidity
- Improving timeliness
  - Using automated case finding for "real time" reporting
  - Automated update of treatment as it is completed.

Automation possibilities:
Increasing the value of registries

- Increasing the clinical and/or public health value of cancer registries by:
  - Including new data that represent important clinical information not currently collected
    - Available electronically
    - Without adding work to the registry staff

An example of the clinical need for new data

There are 10.5 million ***cancer survivors*** (in the US)
estimates reaching 20 million by 2015

- Survival time is no longer sufficient as the sole outcome measure
  - Survivors have multiple recurrences and treatment
  - Registries do not capture this information
    - Precluding the ability to monitor possible complications from chemo
    - Or identifying new complications of treatment over time

What might be done?

An example

The Automated Cancer Extraction (ACE©)

A software system for
using electronic data and automation
to support cancer registration[1]

[1]Software development supported in part by:
Centers for Disease Control & Prevention National Cancer Registry Program
(Modeling Electronic Registry Project (MERP)) and
National Cancer Institute

ACE© Goals

- Enhance efficiency & reduce costs
- Enhance the quality and depth of data
  - Automate capture & upload data
  - Capture increased data details for clinical relevance

Underlying Principles for ACE©

- Uses standardized electronic message formats where feasible (HL7)
- Standardized data formats where feasible (ICD-9 (10), Snomed)
- EMR Platform independent but compatible
- Extensible
- Flexible modular design able to integrate multiple electronic message and data formats

ACE©- What does it do?

Process manager of electronic data

- **Parses** the data into meaningful information
- **Screens** each message or report for cancer relevance based on user-defined text strings and codes
  - to identify new cancers
  - to provide follow up information on known cancer cases
  - to ***automatically*** capture detailed information on treatment
- **Links all reports** at the patient level and stores the data longitudinally

ACE© - Added Value

Automation of data capture & consolidation permits registrars to:

- Evaluate linked information from many sources ***simultaneously*** to inform decision-making
- Maintain source information in a readily accessible system for QA, documentation and for audit purposes
- Consolidated clinical data repository for longitudinal source data for research projects

Electronic Data Sources Feeding ACE ©

- ***Pathology reports***
  - Surgical (HL7 synoptics)
  - Clinical (laboratory tests) (Excel)
- ***Other HL7 reports***
  - Discharge summaries
  - Radiology reports
  - Operative reports
  - Clinical/Physician notes- Oncology, Radiation, Endoscopies
- ***Claims (Billing data)***
  - 837 files (+UB92 & CMS 1500- physician data)
- ***Has the capability to receive other formats***
  - Access files
  - Excel Files
  - SQL Server data
  - Word documents (in process)

Automation opportunities: ACE © Example 1

- Using Nontraditional electronic data sources to enhance registry completeness, detail and timeliness
- ***Electronic billing data***

Automation opportunities: ACE © Example 1 (cont.)

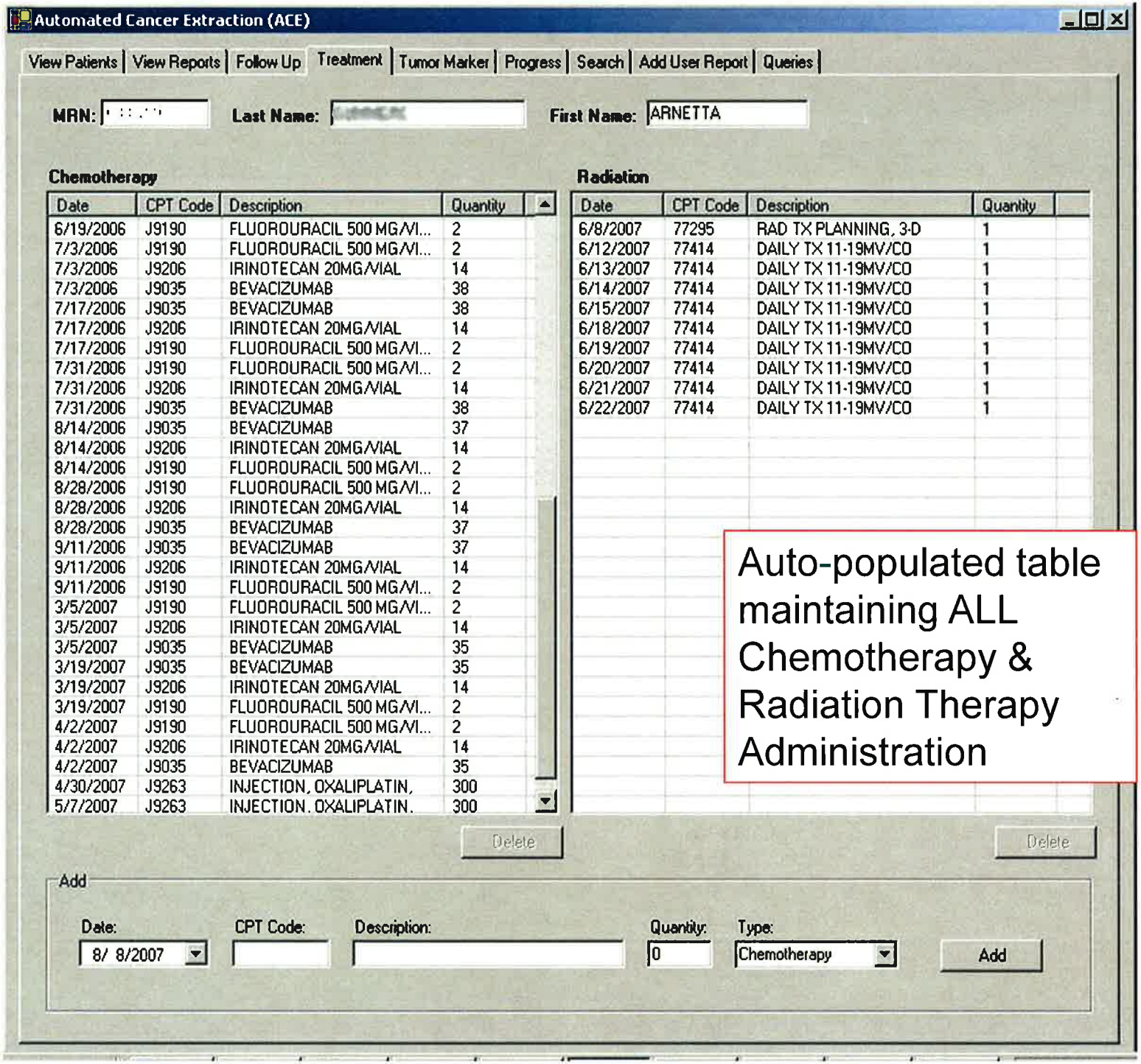
- Electronic billing data
  - Standardized formats for Inpatient and Outpatient data
    - Permit auto-capture & upload of data ***with minimal effort*** to the registrars
    - Provide detailed longitudinal information on:
      - Initial treatment
      - Treatment of recurrent disease
      - Comorbidity information

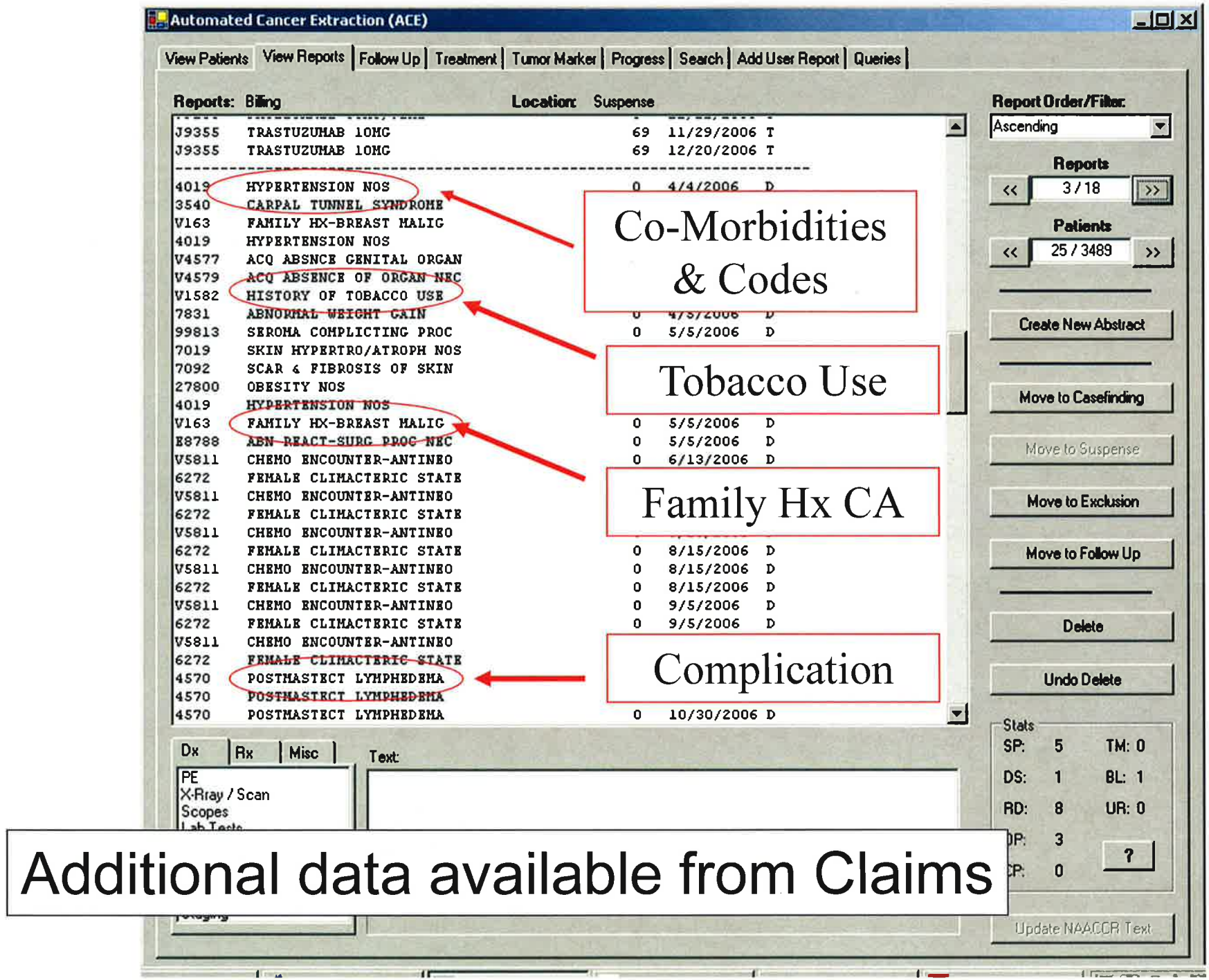
Automation opportunities: ACE © Example 1 (cont.)

Validity of Billing data for treatment

- Studies demonstrate high sensitivity & validity ranging from 88 – 97% [1-4]

1. Warren 2002 (POC/MC: Breast/Colon/Rectum/Ovary)
2. Du 2006 (MC: Breast)
3. Lamont 2005 (CALGB/MC: Breast/Lung)
4. Penberthy 2004 (Br/Lung/CRC/Prostate)


Automated Cancer Extraction (ACE)
View Patients
View Reports
Follow Up
Treatment
Tumor Marker
Progress
Search
Add User Report
Queries
MRN:
Last Name:
First Name: ARNETTA
Chemotherapy
Date | CPT Code | Description | Quantity
6/19/2006 | J9190 | FLUOROURACIL 500 MG/VI... | 2
7/3/2006 | J9190 | FLUOROURACIL 500 MG/VI... | 2
7/3/2006 | J9206 | IRINOTECAN 20MG/VIAL | 14
7/3/2006 | J9035 | BEVACIZUMAB | 38
7/17/2006 | J9035 | BEVACIZUMAB | 38
7/17/2006 | J9206 | IRINOTECAN 20MG/VIAL | 14
7/17/2006 | J9190 | FLUOROURACIL 500 MG/VI... | 2
7/31/2006 | J9190 | FLUOROURACIL 500 MG/VI... | 2
7/31/2006 | J9206 | IRINOTECAN 20MG/VIAL | 14
7/31/2006 | J9035 | BEVACIZUMAB | 38
8/14/2006 | J9035 | BEVACIZUMAB | 37
8/14/2006 | J9206 | IRINOTECAN 20MG/VIAL | 14
8/14/2006 | J9190 | FLUOROURACIL 500 MG/VI... | 2
8/28/2006 | J9190 | FLUOROURACIL 500 MG/VI... | 2
8/28/2006 | J9206 | IRINOTECAN 20MG/VIAL | 14
8/28/2006 | J9035 | BEVACIZUMAB | 37
9/11/2006 | J9035 | BEVACIZUMAB | 37
9/11/2006 | J9206 | IRINOTECAN 20MG/VIAL | 14
9/11/2006 | J9190 | FLUOROURACIL 500 MG/VI... | 2
3/5/2007 | J9190 | FLUOROURACIL 500 MG/VI... | 2
3/5/2007 | J9206 | IRINOTECAN 20MG/VIAL | 14
3/5/2007 | J9035 | BEVACIZUMAB | 35
3/19/2007 | J9035 | BEVACIZUMAB | 35
3/19/2007 | J9206 | IRINOTECAN 20MG/VIAL | 14
3/19/2007 | J9190 | FLUOROURACIL 500 MG/VI... | 2
4/2/2007 | J9190 | FLUOROURACIL 500 MG/VI... | 2
4/2/2007 | J9206 | IRINOTECAN 20MG/VIAL | 14
4/2/2007 | J9035 | BEVACIZUMAB | 35
4/30/2007 | J9263 | INJECTION, OXALIPLATIN, | 300
5/7/2007 | J9263 | INJECTION, OXALIPLATIN, | 300
Delete
Radiation
Date | CPT Code | Description | Quantity
6/8/2007 | 77295 | RAD TX PLANNING, 3-D | 1
6/12/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/13/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/14/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/15/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/18/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/19/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/20/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/21/2007 | 77414 | DAILY TX 11-19MV/CO | 1
6/22/2007 | 77414 | DAILY TX 11-19MV/CO | 1
Delete
Auto-populated table maintaining ALL Chemotherapy & Radiation Therapy Administration
Add
Date: 8/ 8/2007
CPT Code:
Description:
Quantity: 0
Type: Chemotherapy
Add


Automated Cancer Extraction (ACE)
View Patients
View Reports
Follow Up
Treatment
Tumor Marker
Progress
Search
Add User Report
Queries
Reports: Billing
Location: Suspense
J9355 TRASTUZUMAB 10MG 69 11/29/2006 T
J9355 TRASTUZUMAB 10MG 69 12/20/2006 T
4019 HYPERTENSION NOS 0 4/4/2006 D
3540 CARPAL TUNNEL SYNDROME
V163 FAMILY HX-BREAST MALIG
4019 HYPERTENSION NOS
V4577 ACQ ABSNCE GENITAL ORGAN
V4579 ACQ ABSENCE OF ORGAN NEC
V1582 HISTORY OF TOBACCO USE
7831 ABNORMAL WEIGHT GAIN
99813 SEROMA COMPLICTING PROC 0 5/5/2006 D
7019 SKIN HYPERTRO/ATROPH NOS
7092 SCAR & FIBROSIS OF SKIN
27800 OBESITY NOS
4019 HYPERTENSION NOS
V163 FAMILY HX-BREAST MALIG 0 5/5/2006 D
E8788 ABN REACT-SURG PROC NEC 0 5/5/2006 D
V5811 CHEMO ENCOUNTER-ANTINEO 0 6/13/2006 D
6272 FEMALE CLIMACTERIC STATE
V5811 CHEMO ENCOUNTER-ANTINEO
6272 FEMALE CLIMACTERIC STATE
V5811 CHEMO ENCOUNTER-ANTINEO
6272 FEMALE CLIMACTERIC STATE 0 8/15/2006 D
V5811 CHEMO ENCOUNTER-ANTINEO 0 8/15/2006 D
6272 FEMALE CLIMACTERIC STATE 0 8/15/2006 D
V5811 CHEMO ENCOUNTER-ANTINEO 0 9/5/2006 D
6272 FEMALE CLIMACTERIC STATE 0 9/5/2006 D
V5811 CHEMO ENCOUNTER-ANTINEO
6272 FEMALE CLIMACTERIC STATE
4570 POSTMASTECT LYMPHEDEMA
4570 POSTMASTECT LYMPHEDEMA
4570 POSTMASTECT LYMPHEDEMA 0 10/30/2006 D
Co-Morbidities & Codes
Tobacco Use
Family Hx CA
Complication
Report Order/Filter:
Ascending
Reports
<< 3 / 18 >>
Patients
<< 25 / 3489 >>
Create New Abstract
Move to Casefinding
Move to Suspense
Move to Exclusion
Move to Follow Up
Delete
Undo Delete
Stats
SP: 5 TM: 0
DS: 1 BL: 1
RD: 8 UR: 0
3
0
?
Update NAACCR Text
Dx
Rx
Misc
Text:
PE
X-Rray / Scan
Scopes
Additional data available from Claims

Automation opportunities: ACE © Example 2

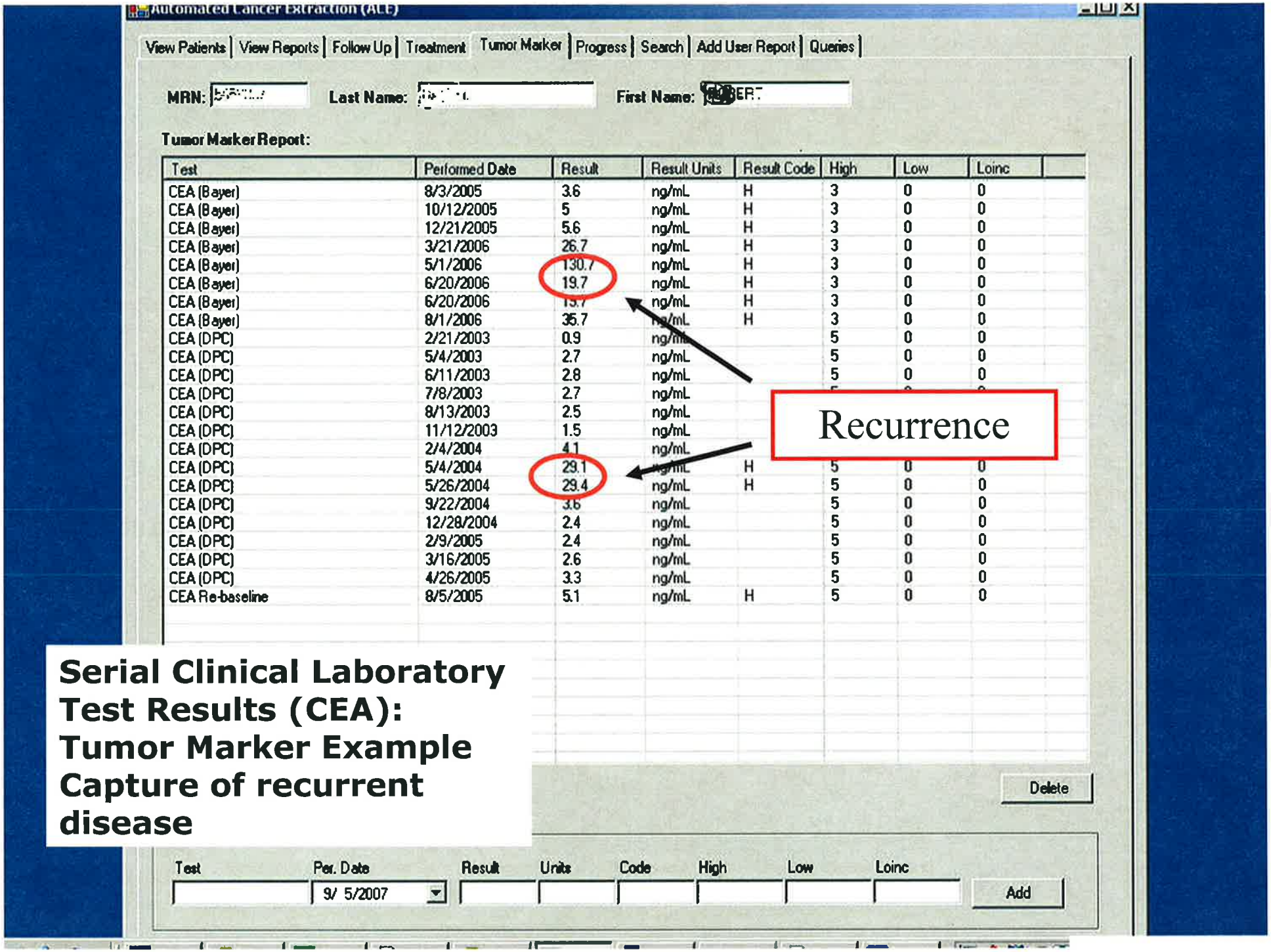
- Using Nontraditional electronic data sources to enhance registry completeness, detail and timeliness

- Clinical laboratory test results
  - Hematologic disease (WBCs, Differential, HgB etc)
  - Serial markers of disease status

Automation opportunities: ACE © Example 2 (cont.)

Clinical Laboratory Test Results:

- Markers for comorbidity
  - Renal function
  - Liver disease
  - Compromised immune system
  - Clinical relevance
  - Implications for treatment choices
- Longitudinal "real time" identification of recurrence using Tumor Marker values



Serial Clinical Laboratory Test Results (CEA): Tumor Marker Example Capture of recurrent disease

Automation opportunities: ACE © Example 3

Auto-population of abstract fields:

- Demographic information from multiple sources
- Treatment information from CPT and ICD-9 procedure codes
- Comorbidity information from billing data
- Follow up data including:
  - Recurrence
  - Treatment for recurrence
- Cancer histopathology data from surgical pathology synoptic reports

ACE Automation: The Need for Checks and Balances

- Review and acceptance by a registrar for critical fields
  - In the cancer abstract
  - Follow up to ascertain recurrence vs second primary cancer

Automation possibilities: Validity and Efficiency of ACE ©

- Can automation make us more efficient?
- Can the data be valid and reliable?

Efficiency: ACE Message Auto-Processing Capacity, 2006

Pathology Reports			
	Processed	27,285	
	Auto-Filtered	***21,480***	***78.72%***
	Reviewed	5,805	
Other Reports			
	Processed	63,313	
	Auto-Filtered	***53,570***	***84.60%***
	Reviewed	9,743	
Billing Messages			
	Processed	116,058	
	Auto-Filtered	***101,821***	***87.70%***
	Reviewed	14,237	

ACE© Efficiency Study Results

- Screening & entering into Suspense of ~2500 patients based on surgical pathology reports
  - Manual: approximately 12 hours
  - ACE: 1.5 hours
  - ACE Identified 9 accessionable cases ***missed manually***
  - Eliminated 11 typing transcription errors

- Reduced average abstraction time from 1.25 hours/case to 0.5 hours per case

Measures of Accuracy for Casefinding Automated vs. Manual Process

Based on Pathology reports only July 2006

	ACE	Manual Process
Sensitivity	**100%**	**77%**
Specificity	**99%**	**99%**
PVP	**79%**	**87%**

Measures of Accuracy/ Effort: False Positives & Overall Predictive Value Positive Rate

- Total "False Positives" 751 Patients in 2006
  - ACE reported probable cases but not accessionable
- "False Positives" Cancer but not reportable= 477 (63%)
  - 449 Non-reportable skin or other cancers
  - 13 History of cancer/in remission
  - 15 cancer history not previously reported

Measures of Accuracy/ Effort: False Positives & Overall Predictive Value Positive Rate

- True "false positives" (not cancer) = 274
  - 135 wi incidental positive cancer "term"
  - 27 wi positive family history of cancer (corrected)
  - 107 Rule out cancer (ICD-9 Dx)
  - 5 Typos in report

- Predictive Value Positive for all data ACE data sources =3438/(274 +3438) (93%)

In summary

- Cancer Registration must increase its value by:
  - Increasing the clinical relevance of cancer data
  - Enhancing data completeness and detail
  - Maintaining or improving efficiency of data collection while:
    - Reducing cost
    - Improving timeliness

In summary

- The increasing availability of electronic data provides an opportunity to meet those challenges
- Judicious use of electronic data & automation coupled with ongoing human review offers significant advantages to current cancer registration processes

Thank you!

For an electronic copy

Email: lpenbert@vcu.edu

- Extra slides

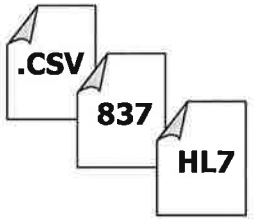
ACE Processes for scanning electronic data for automated capture and upload to registry.

Electronic Data

837 Files
- Inpatient and Outpatient transmission to insurers

HL7 Messages
- Discharge Summary
- Radiology
- Operative Report
- Surgical Pathology

CSV Files
- Clinical Pathology
- Tumor Markers



ACE Process Manager

Cancer relevant term/code?

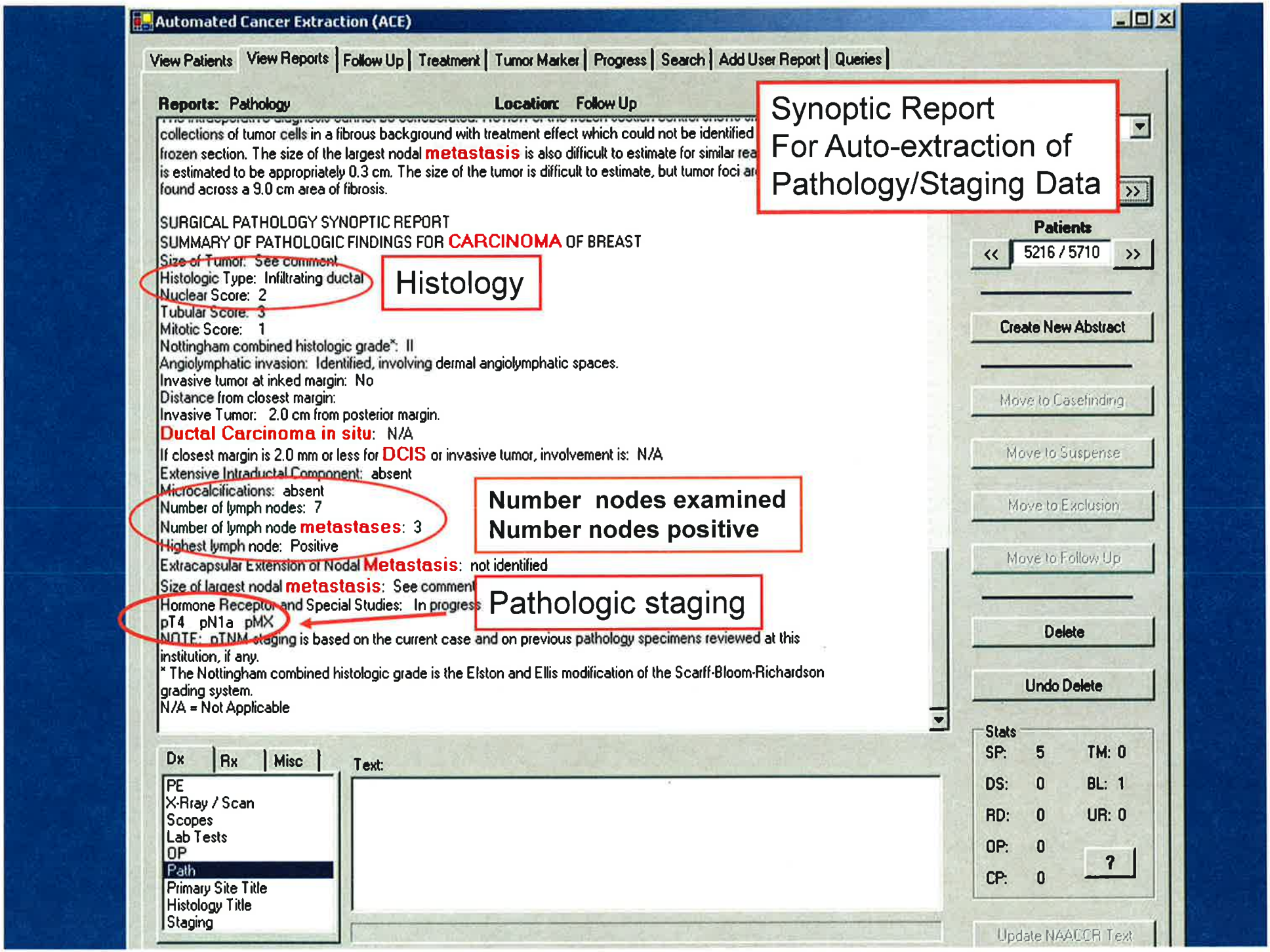
YES
- Match with registry for Existing case for FU or
- Identify as possible new case

NO
- Auto-filter & store in data warehouse

Cancer or Rx term:
Presented to registrar for review/acceptance

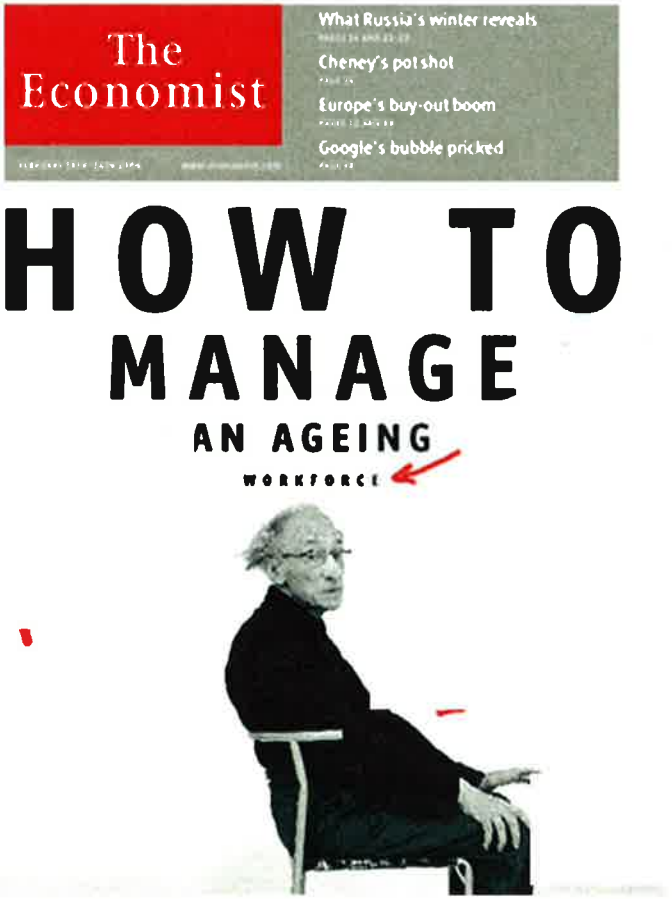
Existing case:
Follow Up info auto -uploaded

Registrar:
- Deletes non-reportable case
- Automated Suspense entry
- Automated abstract initiated
- Treatment detail captured


Automated Cancer Extraction (ACE)
View Patients
View Reports
Follow Up
Treatment
Tumor Marker
Progress
Search
Add User Report
Queries
Reports: Pathology
Location: Follow Up
Synoptic Report
For Auto-extraction of
Pathology/Staging Data
collections of tumor cells in a fibrous background with treatment effect which could not be identified
frozen section. The size of the largest nodal metastasis is also difficult to estimate for similar rea
is estimated to be appropriately 0.3 cm. The size of the tumor is difficult to estimate, but tumor foci ar
found across a 9.0 cm area of fibrosis.
SURGICAL PATHOLOGY SYNOPTIC REPORT
SUMMARY OF PATHOLOGIC FINDINGS FOR CARCINOMA OF BREAST
Size of Tumor: See comment
Histologic Type: Infiltrating ductal
Nuclear Score: 2
Tubular Score: 3
Mitotic Score: 1
Nottingham combined histologic grade*: II
Angiolymphatic invasion: Identified, involving dermal angiolymphatic spaces.
Invasive tumor at inked margin: No
Distance from closest margin:
Invasive Tumor: 2.0 cm from posterior margin.
Ductal Carcinoma in situ: N/A
If closest margin is 2.0 mm or less for DCIS or invasive tumor, involvement is: N/A
Extensive Intraductal Component: absent
Microcalcifications: absent
Number of lymph nodes: 7
Number of lymph node metastases: 3
Highest lymph node: Positive
Extracapsular Extension of Nodal Metastasis: not identified
Size of largest nodal metastasis: See comment
Hormone Receptor and Special Studies: In progress
pT4 pN1a pMX
NOTE: pTNM staging is based on the current case and on previous pathology specimens reviewed at this institution, if any.
* The Nottingham combined histologic grade is the Elston and Ellis modification of the Scarff-Bloom-Richardson grading system.
N/A = Not Applicable
Histology
Number nodes examined
Number nodes positive
Pathologic staging
Patients
<< 5216 / 5710 >>
Create New Abstract
Move to Casefinding
Move to Suspense
Move to Exclusion
Move to Follow Up
Delete
Undo Delete
Stats
SP: 5 TM: 0
DS: 0 BL: 1
RD: 0 UR: 0
OP: 0
CP: 0
?
Update NAACCR Text
Dx
Rx
Misc
PE
X-Rray / Scan
Scopes
Lab Tests
OP
Path
Primary Site Title
Histology Title
Staging
Text:

Challenge 5: ***Limited Skilled Workforce***

- Aging work force
- Limited recruitment of new certified staff
  - primarily US problem
- Tedious routines and extensive clerical chores
  - Reduces efficiency
  - Difficult to maintain qualified personnel at low skill levels



Automation possibilities: Extending a limited workforce

Automation stretches resources by:

- Reducing clerical tasks to reduce boredom
- Likely more appealing to the computer literate generation unwilling to perform clerical functions required in traditional registrations processes